Lecture 3: Mixed quantum states and channels

Mark M. Wilde*

In general, we may not know for certain whether we possess a particular quantum state. Instead, we may only have a probabilistic description of an ensemble of quantum states. This lecture re-establishes the postulates of the quantum theory so that they incorporate a lack of complete information about a quantum system. The density operator formalism is a powerful mathematical tool for describing this scenario. This lecture also establishes how to model the noisy evolution of a quantum system, and we explore models of noisy quantum channels that are the analogs of noisy classical channels.

You might have noticed that the development in the first lecture relied on the premise that the possessor of a quantum system has perfect knowledge of the state of a given system. For instance, we assumed that Alice knows that she possesses a qubit in the state $|\psi\rangle$ where

$$|\psi\rangle = \alpha|0\rangle + \beta|1\rangle. \tag{1}$$

Also, we assumed that Alice and Bob might know that they share an ebit $|\Phi^+\rangle$. We even assumed perfect knowledge of a unitary evolution or a particular measurement that a possessor of a quantum state may apply to it.

This assumption of perfect, definite knowledge of a quantum state is a difficult one to justify in practice. In reality, it is difficult to prepare, evolve, or measure a quantum state exactly as we wish. Slight errors may occur in the preparation, evolution, or measurement due to imprecise devices or to coupling with other degrees of freedom outside of the system that we are controlling. An example of such imprecision can occur in the coupling of two photons at a beamsplitter. We may not be able to tune the reflectivity of the beamsplitter exactly or may not have the timing of the arrival of the photons exactly set. The noiseless quantum theory as we presented it in the previous section cannot handle such imprecisions.

In this lecture, we relax the assumption of perfect knowledge of the preparation, evolution, or measurement of quantum states and develop a noisy quantum theory that incorporates an imprecise knowledge of these states. The noisy quantum theory fuses probability theory and the quantum theory into one formalism.

We proceed with the development of the noisy quantum theory in the following order:

1. We first present the density operator formalism, which gives a representation for a noisy, imprecise quantum state.

2. We then discuss the general form of measurements and the effect of them on our description of a noisy quantum state. We specifically discuss the POVM (positive operator-valued measure) formalism that gives a more general way of describing measurements.

*Mark M. Wilde is with the Department of Physics and Astronomy and the Center for Computation and Technology, Louisiana State University, Baton Rouge, Louisiana 70803. Much of the material is from the book preprint "From Classical to Quantum Shannon Theory" available as `arXiv:1106.1445`.

3. We proceed to composite noisy systems, which admit a particular form, and we discuss several possible states of composite noisy systems including product states, separable states, classical-quantum states, entangled states, and arbitrary states.

4. Next, we consider the Kraus representation of a noisy quantum channel, which gives a way to describe noisy evolution, and we discuss important examples of noisy quantum channels.

1 Noisy Quantum States

We generally may not have perfect knowledge of a prepared quantum state. Suppose a third party, Bob, prepares a state for us and only gives us a probabilistic description of it. We may only know that Bob selects the state $|\psi_x\rangle$ with a certain probability $p_X(x)$. Our description of the state is then as an ensemble $\mathcal{E}$ of quantum states where

$$\mathcal{E} \equiv \{p_X(x), |\psi_x\rangle\}_{x\in\mathcal{X}}. \tag{2}$$

In the above, X is a random variable with distribution $p_X(x)$. Each realization x of random variable X belongs to an alphabet $\mathcal{X}$. For our purposes, it is sufficient for us to say that $\mathcal{X} \equiv \{1, \ldots, |\mathcal{X}|\}$. Thus, the realization x merely acts as an index, meaning that the quantum state is $|\psi_x\rangle$ with probability $p_X(x)$. We also assume that each state $|\psi_x\rangle$ is a qudit state that lives on a system of dimension d.

A simple example is the following ensemble:

$$\left\{\left\{\frac{1}{3}, |1\rangle\right\}, \left\{\frac{2}{3}, |3\rangle\right\}\right\}. \tag{3}$$

The states $|1\rangle$ and $|3\rangle$ live on a four-dimensional system with basis states

$$\{|0\rangle, |1\rangle, |2\rangle, |3\rangle\}. \tag{4}$$

The interpretation of this ensemble is that the state is $|1\rangle$ with probability $1/3$ and the state is $|3\rangle$ with probability $2/3$.

1.1 The Density Operator

Suppose now that we have the ability to perform a perfect measurement of a system with ensemble description $\mathcal{E}$ in (2). Let Π_j be the elements of this projective measurement so that $\sum_j \Pi_j = I$, and let J be the random variable that denotes the index j of the measurement outcome. Let us suppose at first, without loss of generality, that the state in the ensemble is $|\psi_x\rangle$ for some $x \in \mathcal{X}$. Then the Born rule of the noiseless quantum theory states that the conditional probability $p_{J|X}(j|x)$ of obtaining measurement result j (given that the state is $|\psi_x\rangle$) is

$$p_{J|X}(j|x) = \langle\psi_x|\Pi_j|\psi_x\rangle, \tag{5}$$

and the post-measurement state is

$$\frac{\Pi_j|\psi_x\rangle}{\sqrt{p_{J|X}(j|x)}}. \tag{6}$$

But, we would also like to know the actual probability $p_J(j)$ of obtaining measurement result j for the ensemble description $\mathcal{E}$. By the *law of total probability*, the unconditional probability $p_J(j)$ is

$$p_J(j) = \sum_{x\in\mathcal{X}} p_{J|X}(j|x)p_X(x) \tag{7}$$
$$= \sum_{x\in\mathcal{X}} \langle\psi_x|\Pi_j|\psi_x\rangle p_X(x). \tag{8}$$

The *trace* $\text{Tr}\{A\}$ of an operator A is

$$\text{Tr}\{A\} \equiv \sum_i \langle i|A|i\rangle, \tag{9}$$

where $|i\rangle$ is some complete, orthonormal basis. (Observe that the trace operation is *linear*.) We can then show the following useful property with the above definition:

$$\text{Tr}\{\Pi_j|\psi_x\rangle\langle\psi_x|\} = \sum_i \langle i|\Pi_j|\psi_x\rangle\langle\psi_x|i\rangle \tag{10}$$
$$= \sum_i \langle\psi_x|i\rangle\langle i|\Pi_j|\psi_x\rangle \tag{11}$$
$$= \langle\psi_x|\left(\sum_i |i\rangle\langle i|\right)\Pi_j|\psi_x\rangle \tag{12}$$
$$= \langle\psi_x|\Pi_j|\psi_x\rangle. \tag{13}$$

The last equality uses the completeness relation $\sum_i |i\rangle\langle i| = I$. Thus, we continue with the development in (8) and show that

$$p_J(j) = \sum_{x\in\mathcal{X}} \text{Tr}\{\Pi_j|\psi_x\rangle\langle\psi_x|\}p_X(x) \tag{14}$$
$$= \text{Tr}\left\{\Pi_j \sum_{x\in\mathcal{X}} p_X(x)|\psi_x\rangle\langle\psi_x|\right\}. \tag{15}$$

We can rewrite the last equation as follows:

$$p_J(j) = \text{Tr}\{\Pi_j\rho\}, \tag{16}$$

where we define the *density operator* ρ as

$$\rho \equiv \sum_{x\in\mathcal{X}} p_X(x)|\psi_x\rangle\langle\psi_x|. \tag{17}$$

The above operator is known as the density operator because it is the quantum analog of a probability density function.

We sometimes refer to the density operator as the *expected density operator* because there is a sense in which we are taking the expectation over all of the states in the ensemble in order to obtain the density operator. We can equivalenty write the density operator as follows:

$$\rho = \mathbb{E}_X\{|\psi_X\rangle\langle\psi_X|\}, \tag{18}$$

where the expectation is with respect to the random variable X. Note that we are careful to use the notation $|\psi_X\rangle$ instead of the notation $|\psi_x\rangle$ for the state inside of the expectation because the state $|\psi_X\rangle$ is a random quantum state, random with respect to a classical random variable X.

Exercise 1 *Suppose the ensemble has a degenerate probability distribution, say $p_X(0) = 1$ and $p_X(x) = 0$ for all $x \neq 0$. What is the density operator of this degenerate ensemble?*

1.1.1 Properties of the Density Operator

What are the properties that a given density operator must satisfy? Let us consider taking the trace of ρ:

$$\begin{aligned} \text{Tr}\{\rho\} &= \text{Tr}\left\{\sum_{x\in\mathcal{X}} p_X(x)|\psi_x\rangle\langle\psi_x|\right\} && (19)\\ &= \sum_{x\in\mathcal{X}} p_X(x)\text{Tr}\{|\psi_x\rangle\langle\psi_x|\} && (20)\\ &= \sum_{x\in\mathcal{X}} p_X(x)\langle\psi_x|\psi_x\rangle && (21)\\ &= \sum_{x\in\mathcal{X}} p_X(x) && (22)\\ &= 1. && (23) \end{aligned}$$

The above development shows that every density operator should have *unit trace* because it arises from an ensemble of quantum states. Every density operator is also *positive*, meaning that

$$\forall|\varphi\rangle: \qquad \langle\varphi|\rho|\varphi\rangle \geq 0. \tag{24}$$

We write $\rho \geq 0$ to indicate that an operator is positive. The proof of positivity of any density operator ρ is as follows:

$$\begin{aligned} \langle\varphi|\rho|\varphi\rangle &= \langle\varphi|\left(\sum_{x\in\mathcal{X}} p_X(x)|\psi_x\rangle\langle\psi_x|\right)|\varphi\rangle && (25)\\ &= \sum_{x\in\mathcal{X}} p_X(x)\langle\varphi|\psi_x\rangle\langle\psi_x|\varphi\rangle && (26)\\ &= \sum_{x\in\mathcal{X}} p_X(x)|\langle\varphi|\psi_x\rangle|^2 \geq 0. && (27) \end{aligned}$$

The inequality follows because each $p_X(x)$ is a probability and is therefore non-negative.

Let us consider taking the conjugate transpose of the density operator ρ:

$$\begin{aligned} \rho^\dagger &= \left(\sum_{x\in\mathcal{X}} p_X(x)|\psi_x\rangle\langle\psi_x|\right)^\dagger && (28)\\ &= \sum_{x\in\mathcal{X}} p_X(x)(|\psi_x\rangle\langle\psi_x|)^\dagger && (29)\\ &= \sum_{x\in\mathcal{X}} p_X(x)|\psi_x\rangle\langle\psi_x| && (30)\\ &= \rho. && (31) \end{aligned}$$

Every density operator is thus a *Hermitian* operator as well because the conjugate transpose of ρ is ρ.

1.1.2 Ensembles and the Density Operator

Every ensemble has a unique density operator, but the opposite does not necessarily hold: every density operator does not correspond to a unique ensemble and could correspond to many ensembles.

Exercise 2 *Show that the ensembles*

$$\{\{1/2, |0\rangle\}, \{1/2, |1\rangle\}\} \tag{32}$$

and

$$\{\{1/2, |+\rangle\}, \{1/2, |-\rangle\}\} \tag{33}$$

have the same density operator.

This last result has profound implications for the predictions of the quantum theory because it is possible for two or more completely different ensembles to have the same probabilities for measurement results. It also has important implications for quantum Shannon theory as well.

By the spectral theorem, it follows that every density operator ρ has a spectral decomposition in terms of its eigenstates $\{|\phi_x\rangle\}_{x\in\{0,\ldots,d-1\}}$ because every ρ is Hermitian:

$$\rho = \sum_{x=0}^{d-1} \lambda_x |\phi_x\rangle\langle\phi_x|, \tag{34}$$

where the coefficients λ_x are the eigenvalues.

Exercise 3 *Show that the coefficients λ_x are probabilities using the facts that $\mathrm{Tr}\{\rho\} = 1$ and $\rho \geq 0$.*

Thus, given any density operator ρ, we can define a "canonical" ensemble $\{\lambda_x, |\phi_x\rangle\}$ corresponding to it. The ensemble arising from the spectral theorem is the most "efficient" ensemble, in a sense, and we will explore this idea more in the lecture on quantum compression (known as Schumacher compression after its inventor).

1.1.3 Density Operator as the State

We can also refer to the density operator as the *state* of a given quantum system because it is possible to use it to calculate all of the predictions of the quantum theory. We can make these calculations without having an ensemble description—all we need is the density operator. The noisy quantum theory also subsumes the noiseless quantum theory because any state $|\psi\rangle$ has a corresponding density operator $|\psi\rangle\langle\psi|$ in the noisy quantum theory, and all calculations with this density operator in the noisy quantum theory give the same results as using the state $|\psi\rangle$ in the noiseless quantum theory. For these reasons, we will say that the *state* of a given quantum system is a density operator.

One of the most important states in the noisy quantum theory is the maximally mixed state π. The maximally mixed state π arises as the density operator of a uniform ensemble of orthogonal states $\left\{\frac{1}{d}, |x\rangle\right\}$, where d is the dimensionality of the Hilbert space. The maximally mixed state π is then equal to

$$\pi = \frac{1}{d}\sum_{x\in\mathcal{X}}|x\rangle\langle x| = \frac{I}{d}. \tag{35}$$

Exercise 4 *Show that π is the density operator of the ensemble that chooses $|0\rangle$, $|1\rangle$, $|+\rangle$, $|-\rangle$ with equal probability.*

The purity $P(\rho)$ of a density operator ρ is equal to

$$P(\rho) \equiv \mathrm{Tr}\left\{\rho^\dagger\rho\right\} = \mathrm{Tr}\left\{\rho^2\right\}. \tag{36}$$

The purity is one particular measure of the noisiness of a quantum state. The purity of a pure state is equal to one, and the purity of a mixed state is strictly less than one.

1.2 Noiseless Evolution of an Ensemble

Quantum states can evolve in a noiseless fashion either according to a unitary operator or a measurement. In this section, we determine the noiseless evolution of an ensemble and its corresponding density operator. (We consider noisy evolution in Section 4.)

1.2.1 Noiseless Unitary Evolution of a Noisy State

We first consider noiseless evolution according to some unitary U. Suppose we have the ensemble $\mathcal{E}$ in (2) with density operator ρ. Suppose without loss of generality that the state is $|\psi_x\rangle$. Then the evolution postulate of the noiseless quantum theory gives that the state after the unitary evolution is as follows:

$$U|\psi_x\rangle. \tag{37}$$

This result implies that the evolution leads to a new ensemble

$$\mathcal{E}_U \equiv \{p_X(x), U|\psi_x\rangle\}_{x\in\mathcal{X}}. \tag{38}$$

The density operator of the evolved ensemble is

$$\sum_{x\in\mathcal{X}} p_X(x)U|\psi_x\rangle\langle\psi_x|U^\dagger = U\left(\sum_{x\in\mathcal{X}} p_X(x)|\psi_x\rangle\langle\psi_x|\right)U^\dagger \tag{39}$$

$$= U\rho U^\dagger. \tag{40}$$

Thus, the above relation shows that we can keep track of the evolution of the density operator ρ, rather than worrying about keeping track of the evolution of every state in the ensemble $\mathcal{E}$. It suffices to keep track of only the density operator evolution because this operator is sufficient to determine the predictions of the quantum theory.

1.2.2 Noiseless Measurement of a Noisy State

In a similar fashion, we can analyze the result of a measurement on a system with ensemble description $\mathcal{E}$ in (2). Suppose that we perform a projective measurement with projection operators $\{\Pi_j\}_j$ where $\sum_j \Pi_j = I$. Suppose further without loss of generality that the state in the ensemble is $|\psi_x\rangle$. Then the noiseless quantum theory predicts that the probability of obtaining outcome j conditioned on the index x is

$$p_{J|X}(j|x) = \langle\psi_x|\Pi_j|\psi_x\rangle, \tag{41}$$

and the resulting state is

$$\frac{\Pi_j|\psi_x\rangle}{\sqrt{p_{J|X}(j|x)}}. \tag{42}$$

Supposing that we receive outcome j, then we have a new ensemble:

$$\mathcal{E}_j \equiv \left\{p_{X|J}(x|j), \Pi_j|\psi_x\rangle/\sqrt{p_{J|X}(j|x)}\right\}_{x\in\mathcal{X}}. \tag{43}$$

The density operator for this ensemble is

$$\begin{aligned}
&\sum_{x\in\mathcal{X}} p_{X|J}(x|j)\frac{\Pi_j|\psi_x\rangle\langle\psi_x|\Pi_j}{p_{J|X}(j|x)} \\
&= \Pi_j\left(\sum_{x\in\mathcal{X}} \frac{p_{X|J}(x|j)}{p_{J|X}(j|x)}|\psi_x\rangle\langle\psi_x|\right)\Pi_j && (44)\\
&= \Pi_j\left(\sum_{x\in\mathcal{X}} \frac{p_{J|X}(j|x)p_X(x)}{p_{J|X}(j|x)p_J(j)}|\psi_x\rangle\langle\psi_x|\right)\Pi_j && (45)\\
&= \frac{\Pi_j\left(\sum_{x\in\mathcal{X}} p_X(x)|\psi_x\rangle\langle\psi_x|\right)\Pi_j}{p_J(j)} && (46)\\
&= \frac{\Pi_j\rho\Pi_j}{p_J(j)}. && (47)
\end{aligned}$$

The second equality follows from applying the Bayes rule:

$$p_{X|J}(x|j) = p_{J|X}(j|x)p_X(x)/p_J(j). \tag{48}$$

The above expression gives the evolution of the density operator under a measurement. We can again employ the law of total probability to compute that $p_J(j)$ is

$$\begin{aligned}
p_J(j) &= \sum_{x\in\mathcal{X}} p_{J|X}(j|x)p_X(x) && (49)\\
&= \sum_{x\in\mathcal{X}} p_X(x)\langle\psi_x|\Pi_j|\psi_x\rangle && (50)\\
&= \sum_{x\in\mathcal{X}} p_X(x)\text{Tr}\{|\psi_x\rangle\langle\psi_x|\Pi_j\} && (51)\\
&= \text{Tr}\left\{\sum_{x\in\mathcal{X}} p_X(x)|\psi_x\rangle\langle\psi_x|\Pi_j\right\} && (52)\\
&= \text{Tr}\{\rho\Pi_j\}. && (53)
\end{aligned}$$

We can think of $\text{Tr}\{\rho\Pi_j\}$ intuitively as the area of the shadow of ρ onto the space that the projector Π_j projects.

2 Measurement in the Noisy Quantum Theory

We have described measurement in the quantum theory using a set of projectors that form a resolution of the identity. For example, the set $\{\Pi_j\}_j$ of projectors that satisfy the condition $\sum_j \Pi_j = I$ form a valid von Neumann quantum measurement. A projective measurement is not the most general measurement that we can perform on a quantum system (though it is certainly one valid type of quantum measurement).

The most general quantum measurement consists of a set of measurement operators $\{M_j\}_j$ that satisfy the following completeness condition:

$$\sum_j M_j^\dagger M_j = I. \tag{54}$$

Suppose that we have a pure state $|\psi\rangle$. Given a set of measurement operators of the above form, the probability for obtaining outcome j is

$$p(j) \equiv \langle\psi|M_j^\dagger M_j|\psi\rangle, \tag{55}$$

and the post-measurement state when we receive outcome j is

$$\frac{M_j|\psi\rangle}{\sqrt{p(j)}}. \tag{56}$$

Suppose that we instead have an ensemble $\{p_X(x), |\psi_x\rangle\}$ with density operator ρ. We can carry out the same analysis in (47) to show that the post-measurement state when we measure result j is

$$\frac{M_j \rho M_j^\dagger}{p(j)} \tag{57}$$

where the probability $p(j)$ for obtaining outcome j is

$$p(j) \equiv \text{Tr}\left\{M_j^\dagger M_j \rho\right\}. \tag{58}$$

2.1 POVM Formalism

Sometimes, we simply may not care about the post-measurement state of a quantum measurement, but instead we only care about the probability for obtaining a particular outcome. This situation arises in the transmission of classical data over a quantum channel. In this situation, we are merely concerned with minimizing the error probabilities of the classical transmission. The receiver does not care about the post-measurement state because he no longer needs it in the quantum information processing protocol.

We can specify a measurement of this sort by some set $\{\Lambda_j\}_j$ of operators that satisfy positivity and completeness:

$$\Lambda_j \geq 0, \tag{59}$$

$$\sum_j \Lambda_j = I. \tag{60}$$

The set $\{\Lambda_j\}_j$ of operators is a positive operator-valued measure (POVM). The probability for obtaining outcome j is

$$\langle\psi|\Lambda_j|\psi\rangle, \tag{61}$$

if the state is some pure state $|\psi\rangle$. The probability for obtaining outcome j is

$$\mathrm{Tr}\{\Lambda_j\rho\}, \tag{62}$$

if the state is in a mixed state described by some density operator ρ.

Exercise 5 *Consider the following five "Chrysler" states:*

$$|e_k\rangle \equiv \cos\left(\frac{2\pi k}{5}\right)|0\rangle + \sin\left(\frac{2\pi k}{5}\right)|1\rangle, \tag{63}$$

where $k \in \{0, \dots, 4\}$. These states are the "Chrysler" states because they form a pentagon on the XZ-plane of the Bloch sphere. Show that the following set of states form a valid POVM:

$$\left\{\frac{2}{5}|e_k\rangle\langle e_k|\right\}. \tag{64}$$

Exercise 6 *Suppose we have an ensemble $\{p(x), \rho_x\}$ of density operators and a POVM with elements $\{\Lambda_x\}$ that should identify the states ρ_x with high probability, i.e., we would like $Tr\{\rho_x\Lambda_x\}$ to be as high as possible. The expected success probability of the POVM is then*

$$\sum_x p(x)\, Tr\{\rho_x\Lambda_x\}. \tag{65}$$

Suppose that there exists some operator τ such that

$$\tau \geq p(x)\rho_x, \tag{66}$$

where the condition $\tau \geq p(x)\rho_x$ is the same as $\tau - p(x)\rho_x \geq 0$ (the operator $\tau - p(x)\rho_x$ is a positive operator). Show that $Tr\{\tau\}$ is an upper bound on the expected success probability of the POVM. After doing so, consider the case of encoding n bits into a d-dimensional subspace. By choosing states uniformly at random (in the case of the ensemble $\{2^{-n}, \rho_i\}_{i\in\{0,1\}^n}$), show that the expected success probability is bounded above by $d\,2^{-n}$. Thus, it is not possible to store more than n classical bits in n qubits and have a perfect success probability of retrieval (this is a simplified version of the Holevo bound).

3 Composite Noisy Quantum Systems

We are again interested in the behavior of two or more quantum systems when we join them together. Some of the most exotic, truly "quantum" behavior occurs in joint quantum systems, and we observe a marked departure from the classical world.

3.1 Independent Ensembles

Let us first suppose that we have two independent ensembles for quantum systems A and B. The first quantum system belongs to Alice and the second quantum system belongs to Bob, and they may or may not be spatially separated. Let $\{p_X(x), |\psi_x\rangle\}$ be the ensemble for the system A and let $\{p_Y(y), |\phi_y\rangle\}$ be the ensemble for the system B. Suppose for now that the state on system A is $|\psi_x\rangle$ for some x and the state on system B is $|\phi_y\rangle$ for some y. Then, using the composite system postulate of the noiseless quantum theory, the joint state for a given x and y is $|\psi_x\rangle \otimes |\phi_y\rangle$. The density operator for the joint quantum system is the expectation of the states $|\psi_x\rangle \otimes |\phi_y\rangle$ with respect to the random variables X and Y that describe the individual ensembles:

$$\mathbb{E}_{X,Y}\{(|\psi_X\rangle \otimes |\phi_Y\rangle)(\langle\psi_X| \otimes \langle\phi_Y|)\}. \tag{67}$$

The above expression is equal to the following one:

$$\mathbb{E}_{X,Y}\{|\psi_X\rangle\langle\psi_X| \otimes |\phi_Y\rangle\langle\phi_Y|\}, \tag{68}$$

because $(|\psi_x\rangle \otimes |\phi_y\rangle)(\langle\psi_x| \otimes \langle\phi_y|) = |\psi_x\rangle\langle\psi_x| \otimes |\phi_y\rangle\langle\phi_y|$. We then explicitly write out the expectation as a sum over probabilities:

$$\sum_{x,y} p_X(x)p_Y(y)|\psi_x\rangle\langle\psi_x| \otimes |\phi_y\rangle\langle\phi_y|. \tag{69}$$

We can distribute the probabilities and the sum because the tensor product obeys a distributive property:

$$\sum_x p_X(x)|\psi_x\rangle\langle\psi_x| \otimes \sum_y p_Y(y)|\phi_y\rangle\langle\phi_y|. \tag{70}$$

The density operator for this ensemble admits the following simple form:

$$\rho \otimes \sigma, \tag{71}$$

where ρ is the density operator of the X ensemble and σ is the density operator of the Y ensemble. We can say that Alice's local density operator is ρ and Bob's local density operator is σ. We call a density operator of the above form a *product state*. We should expect the density operator to factorize as it does above because we assumed that the ensembles are independent. There is nothing much that distinguishes this situation from the classical world, except for the fact that the states in each respective ensemble may be non-orthogonal to other states in the same ensemble. But even here, there is some equivalent description of each ensemble in terms of an orthonormal basis so that there is really no difference between this description and a joint probability distribution that factors as two independent distributions.

3.2 Separable States

Let us now consider two systems A and B whose corresponding ensembles are correlated. We describe this correlated ensemble as the joint ensemble

$$\{p_X(x), |\psi_x\rangle \otimes |\phi_x\rangle\}. \tag{72}$$

It is straightforward to verify that the density operator of this correlated ensemble has the following form:

$$\mathbb{E}_X\{(|\psi_X\rangle \otimes |\phi_X\rangle)(\langle\psi_X| \otimes \langle\phi_X|)\} = \sum_x p_X(x)|\psi_x\rangle\langle\psi_x| \otimes |\phi_x\rangle\langle\phi_x|. \tag{73}$$

The above state is a *separable* state. The term "separable" implies that there is no quantum entanglement in the above state, i.e., there is a completely classical procedure that prepares the above state. By ignoring Bob's system, Alice's local density operator is of the form

$$\mathbb{E}_X\{|\psi_X\rangle\langle\psi_X|\} = \sum_x p_X(x)|\psi_x\rangle\langle\psi_x|, \tag{74}$$

and similarly, Bob's local density operator is

$$\mathbb{E}_X\{|\phi_X\rangle\langle\phi_X|\} = \sum_x p_X(x)|\phi_x\rangle\langle\phi_x|. \tag{75}$$

We can generalize this classical preparation procedure one step further. Let us suppose that we first generate a random variable Z according to some distribution $p_Z(z)$. We then generate two other ensembles, conditional on the value of the random variable Z. Let $\{p_{X|Z}(x|z), |\psi_{x,z}\rangle\}$ be the first ensemble and let $\{p_{Y|Z}(y|z), |\phi_{y,z}\rangle\}$ be the second ensemble, where the random variables X and Y are independent when conditioned on Z. Let us label the density operators of the first and second ensembles when conditioned on a particular realization z by ρ_z and σ_z, respectively. It is then straightforward to verify that the density operator of an ensemble created from this classical preparation procedure has the following form:

$$\mathbb{E}_{X,Y,Z}\{(|\psi_{X,Z}\rangle \otimes |\phi_{Y,Z}\rangle)(\langle\psi_{X,Z}| \otimes \langle\phi_{Y,Z}|)\} = \sum_z p_Z(z)\rho_z \otimes \sigma_z. \tag{76}$$

Exercise 7 *By ignoring Bob's system, we can determine Alice's local density operator. Show that*

$$\mathbb{E}_{X,Y,Z}\{|\psi_{X,Z}\rangle\langle\psi_{X,Z}|\} = \sum_z p_Z(z)\rho_z, \tag{77}$$

so that the above expression is the density operator for Alice. It similarly follows that the local density operator for Bob is

$$\mathbb{E}_{X,Y,Z}\{|\phi_{Y,Z}\rangle\langle\phi_{Y,Z}|\} = \sum_z p_Z(z)\sigma_z. \tag{78}$$

The density operator in (76) is the most general form of a separable state because the above procedure is the most general classical preparation procedure (we could generalize further with more ensembles of ensembles, but they would ultimately lead to this form because the set of separable states is a convex set). A bipartite state characterized by a density operator is *entangled* if we cannot write it in the form in (76), as a convex combination of product states.

Exercise 8 *Show that we can always write a separable state as a convex combination of pure product states:*

$$\sum_z p_Z(z)|\phi_z\rangle\langle\phi_z| \otimes |\psi_z\rangle\langle\psi_z|, \tag{79}$$

by manipulating the general form in (76).

3.3 Local Density Operator

3.3.1 A First Example

Consider the entangled Bell state $|\Phi^+\rangle^{AB}$ shared on systems A and B. In the above analyses, we were concerned with determining a local density operator description for both Alice and Bob. Now, we are curious if it is possible to determine such a local density operator description for Alice and Bob with respect to the state $|\Phi^+\rangle^{AB}$.

As a first approach to this issue, recall that the density operator description arises from its usefulness in determining the probabilities of the outcomes of a particular measurement. We say that the density operator is "the state" of the system merely because it is a mathematical representation that allows us to compute the probabilities resulting from a physical measurement. So, if we would like to determine a "local density operator," such a local density operator should predict the result of a local measurement.

Let us consider a local measurement with measurement operators $\{M_m\}_m$ that Alice can perform on her system. The global measurement operators for this local measurement are $\left\{M_m^A \otimes I^B\right\}_m$ because nothing (the identity) happens to Bob's system. The probability of obtaining result m is

$$\begin{aligned}
\left\langle\Phi^+\middle|M_m^A \otimes I^B\middle|\Phi^+\right\rangle &= \frac{1}{2}\sum_{i,j=0}^{1}\langle ii|M_m^A \otimes I^B|jj\rangle && (80)\\
&= \frac{1}{2}\sum_{i,j=0}^{1}\langle i|M_m^A|j\rangle\langle i|j\rangle && (81)\\
&= \frac{1}{2}\left(\langle 0|M_m^A|0\rangle + \langle 1|M_m^A|1\rangle\right) && (82)\\
&= \frac{1}{2}\left(\mathrm{Tr}\left\{M_m^A|0\rangle\langle 0|^A\right\} + \mathrm{Tr}\left\{M_m^A|1\rangle\langle 1|^A\right\}\right) && (83)\\
&= \mathrm{Tr}\left\{M_m^A\frac{1}{2}\left(|0\rangle\langle 0|^A + |1\rangle\langle 1|^A\right)\right\} && (84)\\
&= \mathrm{Tr}\left\{M_m^A\pi^A\right\}. && (85)
\end{aligned}$$

The above steps follow by applying the rules of taking the inner product with respect to tensor product operators. The last line follows by recalling the definition of the maximally mixed state π in (35), where π here is a qubit maximally mixed state.

The above calculation demonstrates that we can predict the result of any local "Alice" measurement using the density operator π. Therefore, it is reasonable to say that Alice's local density operator is π, and we even go as far to say that her *local state* is π. A symmetric calculation shows that Bob's local state is also π.

This result concerning their local density operators may seem strange at first. The following global state gives equivalent predictions for local measurements:

$$\pi^A \otimes \pi^B. \quad (86)$$

Can we then conclude that an equivalent representation of the global state is the above state? Absolutely not. The global state $|\Phi^+\rangle^{AB}$ and the above state give drastically different predictions for global measurements. Exercise 10 below asks you to determine the probabilities for measuring the global operator $Z^A \otimes Z^B$ when the global state is $|\Phi^+\rangle^{AB}$ or $\pi^A \otimes \pi^B$, and the result is that the predictions are dramatically different.

Exercise 9 *Show that the projection operators corresponding to a measurement of the observable* $Z^A \otimes Z^B$ *are as follows:*

$$\Pi_{even} \equiv \frac{1}{2}\big(I^A \otimes I^B + Z^A \otimes Z^B\big) = |00\rangle\langle 00|^{AB} + |11\rangle\langle 11|^{AB}, \tag{87}$$

$$\Pi_{odd} \equiv \frac{1}{2}\big(I^A \otimes I^B - Z^A \otimes Z^B\big) = |01\rangle\langle 01|^{AB} + |10\rangle\langle 10|^{AB}. \tag{88}$$

This measurement is a parity measurement, where measurement operator Π_{even} *coherently measures even parity and measurement operator* Π_{odd} *measures odd parity.*

Exercise 10 *Show that a parity measurement (defined in the previous exercise) of the state* $|\Phi^+\rangle^{AB}$ *returns an even parity result with probability one, and a parity measurement of the state* $\pi^A \otimes \pi^B$ *returns even or odd parity with equal probability. Thus, despite the fact that these states have the same local description, their global behavior is very different. Show that the same is true for the phase parity measurement, given by*

$$\Pi_{X,even} \equiv \frac{1}{2}\big(I^A \otimes I^B + X^A \otimes X^B\big), \tag{89}$$

$$\Pi_{X,odd} \equiv \frac{1}{2}\big(I^A \otimes I^B - X^A \otimes X^B\big). \tag{90}$$

Exercise 11 *Show that the maximally correlated state* $\overline{\Phi}^{AB}$*, where*

$$\overline{\Phi}^{AB} = \frac{1}{2}\Big(|00\rangle\langle 00|^{AB} + |11\rangle\langle 11|^{AB}\Big), \tag{91}$$

gives results for local measurements that are the same as those for the maximally entangled state $|\Phi^+\rangle^{AB}$*. Show that the above parity measurements can distinguish these states.*

3.3.2 Partial Trace

In general, we would like to determine a local density operator that predicts the outcomes of all local measurements without having to resort repeatedly to an analysis like that in (80-85). The general method for determining a local density operator is to employ the *partial trace operation.* For a simple state of the form

$$|x\rangle\langle x| \otimes |y\rangle\langle y|, \tag{92}$$

the partial trace is the following operation:

$$|x\rangle\langle x| \ \mathrm{Tr}\{|y\rangle\langle y|\} = |x\rangle\langle x|, \tag{93}$$

where we "trace out" the second system to determine the local density operator for the first. We define it mathematically as acting on any tensor product of rank-one operators (not necessarily corresponding to a state)

$$|x_1\rangle\langle x_2| \otimes |y_1\rangle\langle y_2|, \tag{94}$$

as follows:

$$\mathrm{Tr}_2\{|x_1\rangle\langle x_2| \otimes |y_1\rangle\langle y_2|\} \equiv |x_1\rangle\langle x_2| \ \mathrm{Tr}\{|y_1\rangle\langle y_2|\} \tag{95}$$

$$= |x_1\rangle\langle x_2| \ \langle y_1|y_2\rangle. \tag{96}$$

The subscript "2" of the trace operation indicates that the partial trace acts on the second system. It is a linear operation, much like the full trace is a linear operation.

Exercise 12 *Show that the partial trace operation is equivalent to*

$$Tr_B\Big\{|x_1\rangle\langle x_2|^A \otimes |y_1\rangle\langle y_2|^B\Big\} = \sum_i \langle i|^B \Big(|x_1\rangle\langle x_2|^A \otimes |y_1\rangle\langle y_2|^B\Big)|i\rangle^B, \tag{97}$$

for some orthonormal basis $\{|i\rangle^B\}$ *on Bob's system.*

The most general density operator on two systems A and B is some operator ρ^{AB} that is positive with unit trace. We can obtain the local density operator ρ^A from ρ^{AB} by tracing out the B system:

$$\rho^A = \mathrm{Tr}_B\{\rho^{AB}\}. \tag{98}$$

In more detail, let us expand an arbitrary density operator ρ^{AB} with an orthonormal basis $\{|i\rangle^A \otimes |j\rangle^B\}_{i,j}$ for the bipartite (two-party) state:

$$\rho^{AB} = \sum_{i,j,k,l} \lambda_{i,j,k,l}(|i\rangle^A \otimes |j\rangle^B)(\langle k|^A \otimes \langle l|^B). \tag{99}$$

The coefficients $\lambda_{i,j,k,l}$ are the matrix elements of ρ^{AB} with respect to the basis $\{|i\rangle^A \otimes |j\rangle^B\}_{i,j}$, and they are subject to the constraint of positivity and unit trace for ρ^{AB} . We can rewrite the above operator as

$$\rho^{AB} = \sum_{i,j,k,l} \lambda_{i,j,k,l}|i\rangle\langle k|^A \otimes |j\rangle\langle l|^B. \tag{100}$$

We can now evaluate the partial trace:

$$\begin{aligned}
\rho^A &= \mathrm{Tr}_B\left\{\sum_{i,j,k,l} \lambda_{i,j,k,l}|i\rangle\langle k|^A \otimes |j\rangle\langle l|^B\right\} && (101)\\
&= \sum_{i,j,k,l} \lambda_{i,j,k,l}\mathrm{Tr}_B\Big\{|i\rangle\langle k|^A \otimes |j\rangle\langle l|^B\Big\} && (102)\\
&= \sum_{i,j,k,l} \lambda_{i,j,k,l}|i\rangle\langle k|^A\mathrm{Tr}\Big\{|j\rangle\langle l|^B\Big\} && (103)\\
&= \sum_{i,j,k,l} \lambda_{i,j,k,l}|i\rangle\langle k|^A\langle j|l\rangle && (104)\\
&= \sum_{i,j,k} \lambda_{i,j,k,j}|i\rangle\langle k|^A && (105)\\
&= \sum_{i,k}\left(\sum_j \lambda_{i,j,k,j}\right)|i\rangle\langle k|^A. && (106)
\end{aligned}$$

The second equality exploits the linearity of the partial trace operation. The last equality explicitly shows how the partial trace operation earns its name—it is equivalent to performing a trace operation over the coefficients corresponding to Bob's system.

The next exercise asks you to verify that the operator ρ^A, as defined by the partial trace, predicts the results of a local measurement accurately and confirms the role of ρ^A as a local density operator.

Exercise 13 *Suppose Alice and Bob share a quantum system in a state described by the density operator* ρ^{AB}*. Consider a local measurement, with measurement operators* $\{M_m\}_m$*, that Alice may perform on her system. The global measurement operators are thus* $\{M_m^A \otimes I^B\}_m$*. Show that the probabilities predicted by the global density operator are the same as those predicted by the local density operator* ρ^A *where* $\rho^A = Tr_B\{\rho^{AB}\}$*:*

$$Tr\{(M_m^A \otimes I^B)\rho^{AB}\} = Tr\{M_m^A \rho^A\}. \tag{107}$$

Thus, the predictions of the global noisy quantum theory are consistent with the predictions of the local noisy quantum theory.

Exercise 14 *Verify that the partial trace of a product state gives one of the density operators in the product state:*

$$Tr_2\{\rho \otimes \sigma\} = \rho. \tag{108}$$

This result is consistent with the observation near (71).

Exercise 15 *Verify that the partial trace of a separable state gives the result in (77):*

$$Tr_2\left\{\sum_z p_Z(z)\rho_z \otimes \sigma_z\right\} = \sum_z p_Z(z)\rho_z. \tag{109}$$

Exercise 16 *Consider the following density operator that is formally analogous to a joint probability distribution* $p_{X,Y}(x,y)$*:*

$$\rho = \sum_{x,y} p_{X,Y}(x,y)|x\rangle\langle x| \otimes |y\rangle\langle y|, \tag{110}$$

where the set of states $\{|x\rangle\}_x$ *and* $\{|y\rangle\}_y$ *each form an orthonormal basis. Show that tracing out the second system is formally analogous to taking the marginal distribution* $p_X(x) = \sum_y p_{X,Y}(x,y)$ *of the joint distribution* $p_{X,Y}(x,y)$*. That is, we are left with a density operator of the form*

$$\sum_x p_X(x)|x\rangle\langle x|. \tag{111}$$

Keep in mind that the partial trace is a generalization of the marginalization because it handles more exotic quantum states besides the above "classical" state.

Exercise 17 *Show that the two partial traces in any order on a bipartite system are equivalent to a full trace:*

$$Tr\{\rho^{AB}\} = Tr_A\{Tr_B\{\rho^{AB}\}\} = Tr_B\{Tr_A\{\rho^{AB}\}\}. \tag{112}$$

Exercise 18 *Verify that Alice's local density operator does not change if Bob performs a unitary operator or a measurement where he does not inform her of the measurement result.*

4 Noisy Evolution

The evolution of a quantum state is never perfect. In this section, we introduce noise as resulting from the loss of information about a quantum system. This loss of information is similar to the lack of information about the preparation of a quantum state, as we have seen in the previous section.

4.1 Noisy Evolution from a Random Unitary

We begin with an example, the quantum bit-flip channel. Suppose that we prepare a quantum state $|\psi\rangle$. For simplicity, let us suppose that we are able to prepare this state perfectly. Suppose that we send this qubit over a quantum bit-flip channel, i.e., the channel applies the X Pauli operator (bit-flip operator) with some probability p and applies the identity operator with probability $1-p$. We can describe the resulting state as the following ensemble:

$$\{\{p, X|\psi\rangle\}, \{1-p, |\psi\rangle\}\}. \tag{113}$$

The density operator of this ensemble is as follows:

$$pX|\psi\rangle\langle\psi|X^\dagger + (1-p)|\psi\rangle\langle\psi|. \tag{114}$$

We now generalize the above example by beginning with an ensemble

$$\{p_X(x), |\psi_x\rangle\}_{x\in\mathcal{X}} \tag{115}$$

with density operator $\rho \equiv \sum_{x\in\mathcal{X}} p_X(x)|\psi_x\rangle\langle\psi_x|$ and apply the bit-flip channel to this ensemble. Given that the input state is $|\psi_x\rangle$, the resulting ensemble is as in (113) with $|\psi\rangle$ replaced by $|\psi_x\rangle$. The overall ensemble is then as follows:

$$\{\{p_X(x)p, X|\psi_x\rangle\}, \{p_X(x)(1-p), |\psi_x\rangle\}\}_{x\in\mathcal{X}}. \tag{116}$$

We can calculate the density operator of the above ensemble:

$$\sum_{x\in\mathcal{X}} p_X(x)pX|\psi_x\rangle\langle\psi_x|X^\dagger + p_X(x)(1-p)|\psi_x\rangle\langle\psi_x|, \tag{117}$$

and simplify the above density operator by employing the definition of ρ:

$$pX\rho X^\dagger + (1-p)\rho. \tag{118}$$

The above density operator is more "mixed" than the original density operator.

4.1.1 Random Unitaries

The generalization of the above discussion is to consider some ensemble of unitaries (a random unitary) $\{p(k), U_k\}$ that we can apply to an ensemble of states $\{p_X(x), |\psi_x\rangle\}_{x\in\mathcal{X}}$. It is straightforward to show that the resulting density operator is

$$\sum_k p(k) U_k \rho U_k^\dagger, \tag{119}$$

where ρ is the density operator of the ensemble of states.

4.2 Noisy Evolution as the Loss of a Measurement Outcome

We can also think about noise as arising from the loss of a measurement outcome. Suppose that we have an ensemble of states $\{p_X(x), |\psi_x\rangle\}_{x\in\mathcal{X}}$ and we perform a measurement with a set $\{M_k\}$ of measurement operators where $\sum_k M_k^\dagger M_k = I$. First let us suppose that we know that the state is $|\psi_x\rangle$. Then the probability of obtaining the measurement outcome k is $p_{K|X}(k|x)$ where

$$p_{K|X}(k|x) = \langle\psi_x|M_k^\dagger M_k|\psi_x\rangle, \tag{120}$$

and the post-measurement state is

$$\frac{M_k|\psi_x\rangle}{\sqrt{p_{K|X}(k|x)}}. \tag{121}$$

Let us now suppose that we lose track of the measurement outcome, or equivalently, someone else measures the system and does not inform us of the measurement outcome. The resulting ensemble description is then

$$\left\{p_{X|K}(x|k)p_K(k), M_k|\psi_x\rangle/\sqrt{p_{K|X}(k|x)}\right\}_{x\in\mathcal{X},k}. \tag{122}$$

The density operator of the ensemble is then

$$\begin{aligned}
&\sum_{x,k} p_{X|K}(x|k)p_K(k)\frac{M_k|\psi_x\rangle\langle\psi_x|M_k^\dagger}{p_{K|X}(k|x)} \\
&= \sum_{x,k} p_{K|X}(k|x)p_X(x)\frac{M_k|\psi_x\rangle\langle\psi_x|M_k^\dagger}{p_{K|X}(k|x)} && (123)\\
&= \sum_{x,k} p_X(x)M_k|\psi_x\rangle\langle\psi_x|M_k^\dagger && (124)\\
&= \sum_k M_k\rho M_k^\dagger. && (125)
\end{aligned}$$

We can thus write this evolution as a noisy map $\mathcal{N}(\rho)$ where

$$\mathcal{N}(\rho) \equiv \sum_k M_k\rho M_k^\dagger. \tag{126}$$

We derived the map in (126) from the perspective of the loss of a measurement outcome, but it in fact represents a general evolution of a density operator, and the operators M_k are known as the *Kraus operators*. We can represent all noisy evolutions in the form (126). The evolution of the density operator ρ is *trace-preserving* because the trace of the resulting density operator has unit trace:

$$\begin{aligned}
\mathrm{Tr}\{\mathcal{N}(\rho)\} &= \mathrm{Tr}\left\{\sum_k M_k\rho M_k^\dagger\right\} && (127)\\
&= \sum_k \mathrm{Tr}\left\{M_k\rho M_k^\dagger\right\} && (128)\\
&= \sum_k \mathrm{Tr}\left\{M_k^\dagger M_k\rho\right\} && (129)
\end{aligned}$$

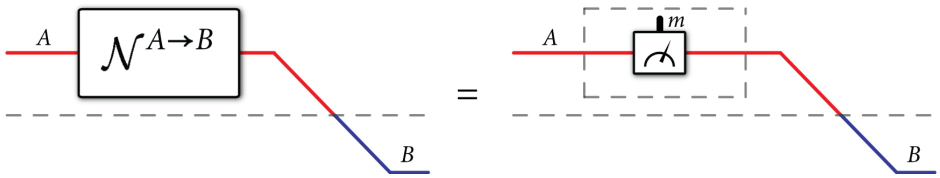

Figure 1: We use the diagram on the left to depict a noisy quantum channel $\mathcal{N}^{A\rightarrow B}$ that takes a quantum system A to a quantum system B. This quantum channel is equivalent to the diagram on the right, where some third party performs a measurement on the input system and does not inform the receiver of the measurement outcome.

$$
\begin{aligned}
&= \mathrm{Tr}\left\{\sum_k M_k^\dagger M_k \rho\right\} && (130) \\
&= \mathrm{Tr}\{\rho\} && (131) \\
&= 1. && (132)
\end{aligned}
$$

There is another important condition that the map $\mathcal{N}(\rho)$ should satisfy: *complete positivity. Positivity* is a special case of complete positivity, and this condition is that the output $\mathcal{N}(\rho)$ is a positive operator whenever the input ρ is a positive operator. Positivity ensures that the noisy evolution produces a quantum state as an output whenever the input is a quantum state. Complete positivity is that the output of the tensor product map $\left(I^k \otimes \mathcal{N}\right)(\sigma)$ for any finite k is a positive operator whenever the input σ is a positive operator (this input operator now lives on a tensor-product Hilbert space). If the input dimension of the noisy map is d, then it is sufficient to consider $k = d$. Complete positivity makes good physical sense because we expect that the action of a noisy map on one system of a quantum state and the identity on the other part of that quantum state should produce as output another quantum state (which is a positive operator). The map $\mathcal{N}(\rho)$ is a completely positive trace-preserving map, and any physical evolution is such a map.

Exercise 19 *Show that the evolution in (126) is positive, i.e., the evolution takes a positive density operator to a positive density operator.*

Exercise 20 *Show that the evolution in (126) is completely positive.*

Exercise 21 *Show that the evolution in (126) is linear:*

$$\mathcal{N}\left(\sum_x p_X(x)\rho_x\right) = \sum_x p_X(x)\mathcal{N}(\rho_x), \tag{133}$$

for any probabilities $p_X(x)$ and density operators ρ_x.

Unitary evolution is a special case of the evolution in (126). We can think of it merely as some measurement where we always know the measurement outcome. That is, it is a measurement with one operator U in the set of measurement operators and it satisfies the completeness condition because $U^\dagger U = I$.

A completely positive trace-preserving map is the mathematical model that we use for a quantum channel in quantum Shannon theory because it represents the most general noisy evolution of a quantum state. This evolution is a generalization of the conditional probability distribution noise model of classical information theory. To see this, suppose that the input density operator ρ is of the following form:

$$\rho = \sum_x p_X(x)|x\rangle\langle x|, \tag{134}$$

where $\{|x\rangle\}$ is some orthonormal basis. We consider a channel $\mathcal{N}$ with the Kraus operators

$$\left\{\sqrt{p_{Y|X}(y|x)}|y\rangle\langle x|\right\}_{x,y}, \tag{135}$$

where $|y\rangle$ and $|x\rangle$ are part of the same basis. Evolution according to this map is then as follows:

$$\begin{aligned}
\mathcal{N}(\rho) &= \sum_{x,y} \sqrt{p_{Y|X}(y|x)}|y\rangle\langle x| \left(\sum_{x'} p_X(x')\left|x'\right\rangle\left\langle x'\right|\right) \sqrt{p_{Y|X}(y|x)}|x\rangle\langle y| && (136)\\
&= \sum_{x,y,x'} p_{Y|X}(y|x) p_X(x')\left|\left\langle x'|x\right\rangle\right|^2 \ |y\rangle\langle y| && (137)\\
&= \sum_{x,y} p_{Y|X}(y|x) p_X(x) \ |y\rangle\langle y| && (138)\\
&= \sum_y \left(\sum_x p_{Y|X}(y|x) p_X(x)\right) \ |y\rangle\langle y|. && (139)
\end{aligned}$$

Thus, the evolution is the same that a noisy classical channel $p_{Y|X}(y|x)$ would enact on a probability distribution $p_X(x)$ by taking it to

$$p_Y(y) = \sum_x p_{Y|X}(y|x) p_X(x) \tag{140}$$

at the output.

4.3 Noisy Evolution from a Unitary Interaction

There is another perspective on quantum noise that is useful to consider. Suppose that a quantum system A begins in the state ρ^A and that there is an environment system E in a pure state $|0\rangle^E$. So the initial state of the joint system AE is

$$\rho^A \otimes |0\rangle\langle 0|^E. \tag{141}$$

Suppose that these two systems interact according to some unitary operator U^{AE} acting on the tensor-product space of A and E. If we are only interested in the state σ^A of the system A after the interaction, then we find it by taking the partial trace over the environment E:

$$\sigma^A = \mathrm{Tr}_E\left\{U^{AE}\left(\rho^A \otimes |0\rangle\langle 0|^E\right)\left(U^{AE}\right)^\dagger\right\}. \tag{142}$$

This evolution is equivalent to that of a completely-positive, trace-preserving map with Kraus operators $\{B_i \equiv \langle i|^E U^{AE}|0\rangle^E\}_i$. This follows easily because we can take the partial trace with respect to an orthonormal basis $\{|i\rangle^E\}$ for the environment:

$$\begin{aligned}
&\mathrm{Tr}_E\Big\{U^{AE}\Big(\rho^A \otimes |0\rangle\langle 0|^E\Big)\big(U^{AE}\big)^\dagger\Big\} \\
&= \sum_i \langle i|^E U^{AE}\Big(\rho^A \otimes |0\rangle\langle 0|^E\Big)\big(U^{AE}\big)^\dagger |i\rangle^E && (143)\\
&= \sum_i \langle i|^E U^{AE}|0\rangle^E \rho^A \langle 0|^E \big(U^{AE}\big)^\dagger |i\rangle^E && (144)\\
&= \sum_i B_i \rho B_i^\dagger. && (145)
\end{aligned}$$

That the operators $\{B_i\}$ are a legitimate set of Kraus operators satisfying $\sum_i B_i^\dagger B_i = I^A$ follows from the unitarity of U^{AE} and the orthonormality of the basis $\{|i\rangle^E\}$:

$$\begin{aligned}
\sum_i B_i^\dagger B_i &= \sum_i \langle 0|^E \big(U^{AE}\big)^\dagger |i\rangle^E \langle i|^E U^{AE}|0\rangle^E && (146)\\
&= \langle 0|^E \big(U^{AE}\big)^\dagger \sum_i |i\rangle\langle i|^E U^{AE}|0\rangle^E && (147)\\
&= \langle 0|^E \big(U^{AE}\big)^\dagger U^{AE}|0\rangle^E && (148)\\
&= \langle 0|^E I^A \otimes I^E|0\rangle^E && (149)\\
&= I^A. && (150)
\end{aligned}$$

4.4 Axiomatic Approach to Noisy Evolution

We can also take an axiomatic approach to understanding noisy quantum evolution, and such an approach leads to the Choi-Kraus representation theorem for quantum channels. It seems reasonable that every evolution of a quantum state should satisfy three properties:

1. It should be linear so that we do not allow for signaling (by a so-called "steering" argument, one could have signaling).

2. It should be completely positive so that it takes quantum states to quantum states (even for systems correlated with the one on which the map is acting).

3. It should be trace preserving (again so that it takes quantum states to quantum states).

The three requirements above lead naturally to Choi's theorem, which states that the map has to take a particular form according to a Kraus decomposition. We now give a proof of Choi's theorem:

Let $|\Gamma\rangle_{RA}$ denote the following vector:

$$|\Gamma\rangle_{RA} \equiv \sum_{i=1}^{d_A} |i\rangle_R \otimes |i\rangle_A, \tag{151}$$

so that it is equal to the maximally entangled state $|\Phi\rangle_{RA}$ times $\sqrt{d_A}$ since

$$|\Phi\rangle_{RA} \equiv \frac{1}{\sqrt{d_A}} \sum_{i=1}^{d_A} |i\rangle_R \otimes |i\rangle_A. \tag{152}$$

Then consider the following Choi matrix of a completely-positive, trace-preserving (CPTP) linear map $\mathcal{N}_{A\to B}$:

$$\mathcal{N}_{A\to B}(|\Gamma\rangle\langle\Gamma|_{RA}) = \sum_{i,j=1}^{d_A} |i\rangle\langle j|_R \otimes \mathcal{N}_{A\to B}(|i\rangle\langle j|). \tag{153}$$

This matrix completely describes the action of the map because it describes the action of it on every operator $|i\rangle\langle j|$, from which we can construct any other operator on which the map acts (it is a large $d_A d_B \times d_A d_B$ matrix with blocks of the form $\mathcal{N}_{A\to B}(|i\rangle\langle j|)$). Also, the above matrix is positive due to the requirement that the map is completely positive. So we can diagonalize $\mathcal{N}_{A\to B}(|\Gamma\rangle\langle\Gamma|_{RA})$ as follows:

$$\mathcal{N}_{A\to B}(|\Gamma\rangle\langle\Gamma|_{RA}) = \sum_{l=1}^{d_A d_B} |\phi_l\rangle\langle\phi_l|_{RB}. \tag{154}$$

(This decomposition does not necessarily have to be such that the vectors $\{|\phi_l\rangle_{RB}\}$ are orthonormal.) Consider by inspecting (153) that

$$(\langle i|_R \otimes I_B)(\mathcal{N}_{A\to B}(|\Gamma\rangle\langle\Gamma|_{RA}))(|j\rangle_R \otimes I_B) = \mathcal{N}_{A\to B}(|i\rangle\langle j|). \tag{155}$$

Now, we can define some operators $\{V_l\}$ by their action on a basis $\{|i\rangle_A\}$ as follows:

$$V_l|i\rangle_A = (\langle i|_R \otimes I_B)|\phi_l\rangle_{RB}. \tag{156}$$

This is related to the observation that any bipartite vector $|\phi_l\rangle_{RB}$ can be written as follows:

$$|\phi_l\rangle_{RB} = I_R \otimes (V_l)_B |\Gamma\rangle_{RB}, \tag{157}$$

for some operator V_l. After making this observation, we finally realize that it is possible to write

$$\begin{aligned}
\mathcal{N}_{A\to B}(|i\rangle\langle j|) &= (\langle i|_R \otimes I_B)(\mathcal{N}_{A\to B}(|\Gamma\rangle\langle\Gamma|_{RA}))(|j\rangle_R \otimes I_B) && (158)\\
&= (\langle i|_R \otimes I_B) \sum_{l=1}^{d_A d_B} |\phi_l\rangle\langle\phi_l|_{RB}(|j\rangle_R \otimes I_B) && (159)\\
&= \sum_{l=1}^{d_A d_B} [(\langle i|_R \otimes I_B)|\phi_l\rangle_{RB}][\langle\phi_l|_{RB}(|j\rangle_R \otimes I_B)] && (160)\\
&= \sum_{l=1}^{d_A d_B} V_l|i\rangle\langle j|_A V_l^\dagger. && (161)
\end{aligned}$$

By linearity of the map $\mathcal{N}_{A\to B}$, it follows that its action on any operator σ can be written as follows:

$$\mathcal{N}_{A\to B}(\sigma) = \sum_{l=1}^{d_A d_B} V_l \sigma V_l^\dagger, \tag{162}$$

since any operator σ can be written as a linear combination of operators in the basis $\{|i\rangle\langle j|\}$.

If the decomposition in (154) is the spectral decomposition, then it follows that the Kraus operators $\{V_l\}$ are orthogonal with respect to the Hilbert-Schmidt inner product:

$$\mathrm{Tr}\Big\{V_l^\dagger V_k\Big\} = \mathrm{Tr}\Big\{V_l^\dagger V_l\Big\}\delta_{l,k}. \tag{163}$$

This follows from the fact that

$$\begin{aligned}
\delta_{l,k}\langle\phi_l|\phi_l\rangle &= \langle\phi_l|\phi_k\rangle && (164)\\
&= \langle\Gamma|_{RB}\Big[I_R \otimes \Big(V_l^\dagger V_k\Big)_B\Big]|\Gamma\rangle_{RB} && (165)\\
&= \sum_{i,j}\langle i|j\rangle\ \langle i|\Big(V_l^\dagger V_k\Big)|j\rangle && (166)\\
&= \sum_i \langle i|\Big(V_l^\dagger V_k\Big)|i\rangle && (167)\\
&= \mathrm{Tr}\Big\{V_l^\dagger V_k\Big\}. && (168)
\end{aligned}$$

The Kraus operators $\{V_l\}$ should satisfy the following condition:

$$\sum_{l=1}^{d_A d_B} V_l^\dagger V_l = I. \tag{169}$$

This follows from the fact that

$$\mathrm{Tr}_B\{\mathcal{N}_{A\to B}(|\Phi\rangle\langle\Phi|_{RA})\} = \frac{1}{d_A} I_R. \tag{170}$$

(That is, when inputting the maximally entangled state, the reduced state on the reference should be the maximally mixed state no matter what the CPTP map $\mathcal{N}_{A\to B}$ is.) By using the fact that

$$(I \otimes C_A)|\Phi\rangle_{RA} = \big(C_R^T \otimes I_A\big)|\Phi\rangle_{RA}, \tag{171}$$

for any operator C, we find that

$$\begin{aligned}
\mathrm{Tr}_B\{\mathcal{N}_{A\to B}(|\Phi\rangle\langle\Phi|_{RA})\} &= \mathrm{Tr}_B\left\{\sum_{l=1}^{d_A d_B}(I_R \otimes V_l)(|\Phi\rangle\langle\Phi|_{RA})\Big(I_R \otimes V_l^\dagger\Big)\right\} && (172)\\
&= \mathrm{Tr}_B\left\{\sum_{l=1}^{d_A d_B}\big(\big(V_l^T\big)_R \otimes I_A\big)(|\Phi\rangle\langle\Phi|_{RA})((V_l^*)_R \otimes I_A)\right\} && (173)\\
&= \frac{1}{d_A}\sum_{l=1}^{d_A d_B} V_l^T V_l^* && (174)\\
&= \frac{1}{d_A} I_R. && (175)
\end{aligned}$$

The result that $\sum_{l=1}^{d_A d_B} V_l^\dagger V_l$ follows because

$$\frac{1}{d_A} I_R = \left(\frac{1}{d_A} I_R\right)^*. \tag{176}$$

4.5 Unique Specification of a Noisy Channel

Consider a given noisy quantum channel $\mathcal{N}$ with Kraus representation

$$\mathcal{N}(\rho) = \sum_j A_j \rho A_j^\dagger. \tag{177}$$

We can also uniquely specify $\mathcal{N}$ by its action on an operator of the form $|i\rangle\langle j|$ where $\{|i\rangle\}$ is some orthonormal basis:

$$N_{ij} \equiv \mathcal{N}(|i\rangle\langle j|). \tag{178}$$

We can figure out how the channel $\mathcal{N}$ would act on any density operator if we know how it acts on $|i\rangle\langle j|$ for all i and j. Thus, two channels $\mathcal{N}$ and $\mathcal{M}$ are equivalent if they have the same effect on all operators of the form $|i\rangle\langle j|$:

$$\mathcal{N} = \mathcal{M} \quad \Leftrightarrow \quad \forall i,j \quad \mathcal{N}(|i\rangle\langle j|) = \mathcal{M}(|i\rangle\langle j|). \tag{179}$$

Let us now consider a maximally entangled qudit state $|\Phi\rangle^{AB}$ where

$$|\Phi\rangle^{AB} = \frac{1}{\sqrt{d}} \sum_{i=0}^{d-1} |i\rangle^A |i\rangle^B, \tag{180}$$

and d is the dimension of each system A and B. The density operator Φ^{AB} of such a state is as follows:

$$\Phi^{AB} = \frac{1}{d} \sum_{i,j=0}^{d-1} |i\rangle\langle j|^A \otimes |i\rangle\langle j|^B. \tag{181}$$

Let us now send the A system of Φ^{AB} through the noisy quantum channel $\mathcal{N}$:

$$\left(\mathcal{N}^A \otimes I^B\right)\left(\Phi^{AB}\right) = \frac{1}{d} \sum_{i,j=0}^{d-1} \mathcal{N}^A(|i\rangle\langle j|^A) \otimes |i\rangle\langle j|^B. \tag{182}$$

The resulting state completely characterizes the noisy channel $\mathcal{N}$ because the following map translates between the state in (182) and the operators N_{ij} in (178):

$$d\langle i'|\left(\mathcal{N}^A \otimes I^B\right)\left(\Phi^{AB}\right)|j'\rangle^B = N_{ij}. \tag{183}$$

Thus, we can completely characterize a noisy map by determining the quantum state resulting from sending half of a maximally entangled state through it, and the following condition is necessary and sufficient for any two noisy channels to be equivalent:

$$\mathcal{N} = \mathcal{M} \quad \Leftrightarrow \quad \left(\mathcal{N}^A \otimes I^B\right)\left(\Phi^{AB}\right) = \left(\mathcal{M}^A \otimes I^B\right)\left(\Phi^{AB}\right). \tag{184}$$

It is equivalent to the condition in (179).

4.6 Concatenation of Noisy Maps

A quantum state may undergo not just one type of noisy evolution—it can of course undergo one noisy quantum channel followed by another noisy quantum channel. Let $\mathcal{N}_1$ denote a first noisy evolution and let $\mathcal{N}_2$ denote a second noisy evolution. Suppose that the Kraus operators of $\mathcal{N}_1$ are $\{A_k\}$ and the Kraus operators of $\mathcal{N}_2$ are $\{B_k\}$. It is straightforward to define the concatenation $\mathcal{N}_2 \circ \mathcal{N}_1$ of these two maps. Consider that the output of the first map is

$$\mathcal{N}_1(\rho) \equiv \sum_k A_k \rho A_k^\dagger, \tag{185}$$

for some input density operator ρ. The output of the concatenation map $\mathcal{N}_2 \circ \mathcal{N}_1$ is then

$$(\mathcal{N}_2 \circ \mathcal{N}_1)(\rho) = \sum_k B_k \mathcal{N}_1(\rho) B_k^\dagger = \sum_{k,k'} B_k A_{k'} \rho A_{k'}^\dagger B_k^\dagger. \tag{186}$$

It is clear that the Kraus operators of the concatenation map are $\{B_k A_{k'}\}_{k,k'}$.

4.7 Important Examples of Noisy Evolution

This section discusses some of the most important examples of noisy evolutions that we will consider in this book. Throughout this book, we will be considering the information-carrying ability of these various channels. They will provide some useful, "hands on" insight into quantum Shannon theory.

4.7.1 Dephasing Channel

We have already given the example of a noisy quantum bit flip channel in Section 4.1. Another important example is a bit flip in the conjugate basis, or equivalently, a *phase flip channel*. This channel acts as follows on any given density operator:

$$\rho \to (1-p)\rho + pZ\rho Z. \tag{187}$$

It is also known as the *dephasing channel*.

For $p = 1/2$, the action of the dephasing channel on a given quantum state is equivalent to the action of measuring the qubit in the computational basis and forgetting the result of the measurement. We make this idea more clear with an example. First, suppose that we have a qubit

$$|\psi\rangle = \alpha|0\rangle + \beta|1\rangle, \tag{188}$$

and we measure it in the computational basis. Then the postulates of quantum theory state that the qubit becomes $|0\rangle$ with probability $|\alpha|^2$ and it becomes $|1\rangle$ with probability $|\beta|^2$. Suppose that we forget the measurement outcome, or alternatively, we do not have access to it. Then our best description of the qubit is with the following ensemble:

$$\left\{\left\{|\alpha|^2, |0\rangle\right\}, \left\{|\beta|^2, |1\rangle\right\}\right\}. \tag{189}$$

The density operator of this ensemble is

$$|\alpha|^2|0\rangle\langle 0| + |\beta|^2|1\rangle\langle 1|. \tag{190}$$

Now let us check if the dephasing channel gives the same behavior as the forgetful measurement above. We can consider the qubit as being an ensemble $\{1, |\psi\rangle\}$, i.e., the state is certain to be $|\psi\rangle$. The density operator of the ensemble is then ρ where

$$\rho = |\alpha|^2 |0\rangle\langle 0| + \alpha\beta^* |0\rangle\langle 1| + \alpha^*\beta |1\rangle\langle 0| + |\beta|^2 |1\rangle\langle 1|. \tag{191}$$

If we act on the density operator ρ with the dephasing channel with $p = 1/2$, then it preserves the density operator with probability 1/2 and phase flips the qubit with probability 1/2:

$$\begin{aligned}
&\frac{1}{2}\rho + \frac{1}{2} Z\rho Z \\
&= \frac{1}{2}\Big(|\alpha|^2 |0\rangle\langle 0| + \alpha\beta^* |0\rangle\langle 1| + \alpha^*\beta |1\rangle\langle 0| + |\beta|^2 |1\rangle\langle 1|\Big) + \\
&\qquad \frac{1}{2}\Big(|\alpha|^2 |0\rangle\langle 0| - \alpha\beta^* |0\rangle\langle 1| - \alpha^*\beta |1\rangle\langle 0| + |\beta|^2 |1\rangle\langle 1|\Big) && (192) \\
&= |\alpha|^2 |0\rangle\langle 0| + |\beta|^2 |1\rangle\langle 1|. && (193)
\end{aligned}$$

The dephasing channel nullifies the off-diagonal terms in the density operator with respect to the computational basis. The resulting density operator description is the same as what we found for the forgetful measurement.

Exercise 22 *Verify that the action of the dephasing channel on the Bloch vector is*

$$\frac{1}{2}(I + r_x X + r_y Y + r_z Z) \rightarrow \frac{1}{2}(I + (1-2p) r_x X + (1-2p) r_y Y + r_z Z), \tag{194}$$

so that the channel preserves any component of the Bloch vector in the Z direction, while shrinking any component in the X or Y direction.

4.7.2 Pauli Channel

A Pauli channel is a generalization of the above dephasing channel and the bit flip channel. It simply applies a random Pauli operator according to a probability distribution. The map for a qubit Pauli channel is

$$\rho \rightarrow \sum_{i,j=0}^{1} p(i,j) Z^i X^j \rho X^j Z^i. \tag{195}$$

The generalization of this channel to qudits is straightforward. We simply replace the Pauli operators with the Heisenberg-Weyl operators. The Pauli qudit channel is

$$\rho \rightarrow \sum_{i,j=0}^{d-1} p(i,j) Z(i) X(j) \rho X^\dagger(j) Z^\dagger(i). \tag{196}$$

These channels are important in the study of quantum key distribution (QKD) because an eavesdropper induces such a channel in a QKD protocol.

Exercise 23 *We can write a Pauli channel as*

$$\rho \rightarrow p_I \rho + p_X X \rho X + p_Y Y \rho Y + p_Z Z \rho Z. \tag{197}$$

Verify that the action of the Pauli channel on the Bloch vector is

$$(r_x, r_y, r_z) \rightarrow \\ ((p_I + p_X - p_Y - p_Z) r_x,\ (p_I + p_Y - p_X - p_Z) r_y,\ (p_I + p_Z - p_X - p_Y) r_z). \tag{198}$$

4.7.3 Depolarizing Channel

The depolarizing channel is a "worst-case scenario" channel. It assumes that we just completely lose the input qubit with some probability, i.e., it replaces the lost qubit with the maximally mixed state. The map for the depolarizing channel is

$$\rho \rightarrow (1-p)\rho + p\pi, \tag{199}$$

where π is the maximally mixed state: $\pi = I/2$.

Most of the time, this channel is too pessimistic. Usually, we can learn something about the physical nature of the channel by some estimation process. We should only consider using the depolarizing channel as a model if we have little to no information about the actual physical channel.

Exercise 24 (Pauli Twirl) *Show that randomly applying the Pauli operators I, X, Y, Z with uniform probability to any density operator gives the maximally mixed state:*

$$\frac{1}{4}\rho + \frac{1}{4} X \rho X + \frac{1}{4} Y \rho Y + \frac{1}{4} Z \rho Z = \pi. \tag{200}$$

(Hint: Represent the density operator as $\rho = (I + r_x X + r_y Y + r_z Z)/2$ and apply the commutation rules of the Pauli operators.) This is known as the "twirling" operation.

Exercise 25 *Show that we can rewrite the depolarizing channel as the following Pauli channel:*

$$\rho \rightarrow (1 - 3p/4)\rho + p\left(\frac{1}{4} X \rho X + \frac{1}{4} Y \rho Y + \frac{1}{4} Z \rho Z\right). \tag{201}$$

Exercise 26 *Show that the action of a depolarizing channel on the Bloch vector is*

$$(r_x, r_y, r_z) \rightarrow ((1-p) r_x,\ (1-p) r_y,\ (1-p) r_z). \tag{202}$$

Thus, it uniformly shrinks the Bloch vector to become closer to the maximally mixed state.

The generalization of the depolarizing channel to qudits is again straightforward. It is the same as the map in (199), with the exception that the density operators ρ and π are qudit density operators.

Exercise 27 (Qudit Twirl) *Show that randomly applying the Heisenberg-Weyl operators*

$$\{X(i)Z(j)\}_{i,j\in\{0,\ldots,d-1\}} \tag{203}$$

with uniform probability to any qudit density operator gives the maximally mixed state π*:*

$$\frac{1}{d^2}\sum_{i,j=0}^{d-1} X(i)Z(j)\rho Z^\dagger(j)X^\dagger(i) = \pi. \tag{204}$$

(Hint: You can do the full calculation, or you can decompose this channel into the composition of two completely dephasing channels where the first is a dephasing in the computational basis and the next is a dephasing in the conjugate basis).

4.7.4 Amplitude Damping Channel

The amplitude damping channel is a first-order approximation to a noisy evolution that occurs in many physical systems ranging from optical systems to chains of spin-1/2 particles to spontaneous emission of a photon from an atom.

In order to motivate this channel, we give a physical interpretation to our computational basis states. Let us think of the $|0\rangle$ state as the ground state of a two-level atom and let us think of the state $|1\rangle$ as the excited state of the atom. Spontaneous emission is a process that tends to decay the atom from its excited state to its ground state, even if the atom is in a superposition of the ground and excited states. Let the parameter γ denote the probability of decay so that $0 \leq \gamma \leq 1$. One Kraus operator that captures the decaying behavior is

$$A_0 = \sqrt{\gamma}|0\rangle\langle 1|. \tag{205}$$

The operator A_0 annihilates the ground state:

$$A_0|0\rangle\langle 0|A_0^\dagger = 0, \tag{206}$$

and it decays the excited state to the ground state:

$$A_0|1\rangle\langle 1|A_0^\dagger = \gamma|0\rangle\langle 0|. \tag{207}$$

The Kraus operator A_0 alone does not specify a physical map because $A_0^\dagger A_0 = \gamma|1\rangle\langle 1|$ (recall that the Kraus operators of any channel should satisfy the condition $\sum_k A_k^\dagger A_k = I$). We can satisfy this condition by choosing another operator A_1 such that

$$\begin{aligned} A_1^\dagger A_1 &= I - A_0^\dagger A_0 && (208)\\ &= |0\rangle\langle 0| + (1-\gamma)|1\rangle\langle 1|. && (209)\end{aligned}$$

The following choice of A_1 satisfies the above condition:

$$A_1 \equiv |0\rangle\langle 0| + \sqrt{1-\gamma}|1\rangle\langle 1|. \tag{210}$$

Thus, the operators A_0 and A_1 are valid Kraus operators for the amplitude damping channel.

Exercise 28 *Consider a single-qubit density operator with the following matrix representation with respect to the computational basis:*

$$\rho = \begin{bmatrix} 1-p & \eta \\ \eta^* & p \end{bmatrix}, \tag{211}$$

where $0 \leq p \leq 1$ and η is some complex number. Show that applying the amplitude damping channel with parameter γ to a qubit with the above density operator gives a density operator with the following matrix representation:

$$\begin{bmatrix} 1-(1-\gamma)p & \sqrt{1-\gamma}\eta \\ \sqrt{1-\gamma}\eta^* & (1-\gamma)p \end{bmatrix}. \tag{212}$$

Exercise 29 *Show that the amplitude damping channel obeys a composition rule. Consider an amplitude damping channel $\mathcal{N}_1$ with transmission parameter $(1-\gamma_1)$ and consider another amplitude damping channel $\mathcal{N}_2$ with transmission parameter $(1-\gamma_2)$. Show that the composition channel $\mathcal{N}_2 \circ \mathcal{N}_1$ is an amplitude damping channel with transmission parameter $(1-\gamma_1)(1-\gamma_2)$. (Note that the transmission parameter is equal to one minus the damping parameter.)*

4.7.5 Erasure Channel

The erasure channel is another important channel in quantum Shannon theory. It admits a simple model and is amenable to relatively straightforward analysis when we later discuss its capacity. The erasure channel can serve as a simplified model of photon loss in optical systems.

We first recall the classical definition of an erasure channel. A classical erasure channel either transmits a bit with some probability $1-\varepsilon$ or replaces it with an erasure symbol e with some probability ε. The output alphabet contains one more symbol than the input alphabet, namely, the erasure symbol e.

The generalization of the classical erasure channel to the quantum world is straightforward. It implements the following map:

$$\rho \to (1-\varepsilon)\rho + \varepsilon|e\rangle\langle e|, \tag{213}$$

where $|e\rangle$ is some state that is not in the input Hilbert space, and thus is orthogonal to it. The output space of the erasure channel is larger than its input space by one dimension. The interpretation of the quantum erasure channel is similar to that for the classical erasure channel. It transmits a qubit with probability $1-\varepsilon$ and "erases" it (replaces it with an orthogonal erasure state) with probability ε.

Exercise 30 *Show that the following operators are the Kraus operators for the quantum erasure channel:*

$$\sqrt{1-\varepsilon}(|0\rangle^B\langle 0|^A + |1\rangle^B\langle 1|^A), \tag{214}$$

$$\sqrt{\varepsilon}|e\rangle^B\langle 0|^A, \tag{215}$$

$$\sqrt{\varepsilon}|e\rangle^B\langle 1|^A. \tag{216}$$

At the receiving end of the channel, a simple measurement can determine whether an erasure has occurred. We perform a measurement with measurement operators $\{\Pi_{\text{in}}, |e\rangle\langle e|\}$, where Π_{in} is the projector onto the input Hilbert space. This measurement has the benefit of detecting no more information than necessary. It merely detects whether an erasure occurs, and thus preserves the quantum information at the input if an erasure does not occur.

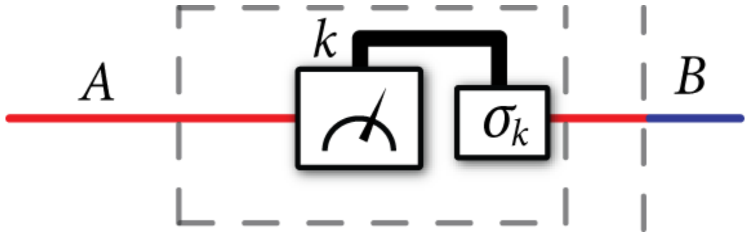

Figure 2: The above figure illustrates the internal workings of a classical-quantum channel. It first measures the input state in some basis $\{|k\rangle\}$ and outputs a quantum state σ_k conditional on the measurement outcome.

4.7.6 Classical-Quantum Channel

A classical-quantum channel is one that first measures the input state in a particular orthonormal basis and outputs a density operator conditional on the result of the measurement. Suppose that the input to the channel is a density operator ρ. Suppose that $\{|k\rangle\}_k$ is an orthonormal basis for the Hilbert space on which the density operator ρ acts. The classical-quantum channel first measures the input state in the basis $\{|k\rangle\}$. Given that the result of the measurement is k, the post measurement state is

$$\frac{|k\rangle\langle k|\rho|k\rangle\langle k|}{\langle k|\rho|k\rangle}. \tag{217}$$

The classical-quantum channel correlates a density operator σ_k with the post-measurement state k:

$$\frac{|k\rangle\langle k|\rho|k\rangle\langle k|}{\langle k|\rho|k\rangle} \otimes \sigma_k. \tag{218}$$

This action leads to an ensemble:

$$\left\{\langle k|\rho|k\rangle, \frac{|k\rangle\langle k|\rho|k\rangle\langle k|}{\langle k|\rho|k\rangle} \otimes \sigma_k\right\}, \tag{219}$$

and the density operator of the ensemble is

$$\sum_k \langle k|\rho|k\rangle \frac{|k\rangle\langle k|\rho|k\rangle\langle k|}{\langle k|\rho|k\rangle} \otimes \sigma_k = \sum_k |k\rangle\langle k|\rho|k\rangle\langle k| \otimes \sigma_k. \tag{220}$$

The channel then only outputs the system on the right (tracing out the first system) so that the resulting channel is as follows:

$$\mathcal{N}(\rho) \equiv \sum_k \langle k|\rho|k\rangle \sigma_k. \tag{221}$$

Figure 2 depicts the behavior of the classical-quantum channel. This channel is a particular kind of entanglement-breaking channel, for reasons that become clear in the next exercise.

Exercise 31 *Show that the classical-quantum channel is an entanglement-breaking channel—i.e., if we input the B system of an entangled state ψ^{AB}, then the resulting state on AB is no longer entangled.*

We can prove a more general structural theorem regarding entanglement-breaking channels by exploiting the observation in the above exercise.

Theorem 32 *An entanglement-breaking channel has a representation with Kraus operators that are unit rank.*

Proof. Consider that the output of an entanglement-breaking channel $\mathcal{N}_{\text{EB}}$ acting on half of a maximally entangled state is as follows:

$$\mathcal{N}_{\text{EB}}^{A\to B'}(\Phi^{BA}) = \sum_z p_Z(z)|\phi_z\rangle\langle\phi_z|^B \otimes |\psi_z\rangle\langle\psi_z|^{B'}. \tag{222}$$

This holds because the output of a channel is a separable state (it "breaks" entanglement), and it is always possible to find a representation of the separable state with pure states (see Exercise 8). Now consider constructing a channel $\mathcal{M}$ with the following unit-rank Kraus operators:

$$A_z \equiv \left\{\sqrt{d\ p_Z(z)}|\psi_z\rangle\langle\phi_z^*|\right\}_z, \tag{223}$$

where $|\phi_z^*\rangle$ is the state $|\phi_z\rangle$ with all of its elements conjugated. We should first verify that these Kraus operators form a valid channel, by checking that $\sum_z A_z^\dagger A_z = I$:

$$\begin{aligned}\sum_z A_z^\dagger A_z &= \sum_z d\ p_Z(z)|\phi_z^*\rangle\langle\psi_z|\psi_z\rangle\langle\phi_z^*| &(224)\\ &= d\ \sum_z p_Z(z)|\phi_z^*\rangle\langle\phi_z^*|. &(225)\end{aligned}$$

Consider that

$$\begin{aligned}\text{Tr}_{B'}\left\{\mathcal{N}_{\text{EB}}^{A\to B'}(\Phi^{BA})\right\} &= \pi^B &(226)\\ &= \text{Tr}_{B'}\left\{\sum_z p_Z(z)|\phi_z\rangle\langle\phi_z|^B \otimes |\psi_z\rangle\langle\psi_z|^{B'}\right\} &(227)\\ &= \sum_z p_Z(z)|\phi_z\rangle\langle\phi_z|^B, &(228)\end{aligned}$$

where π^B is the maximally mixed state. Thus, it follows that $\mathcal{M}$ is a valid quantum channel because

$$\begin{aligned}d\ \sum_z p_Z(z)|\phi_z\rangle\langle\phi_z|^B &= d\ \pi^B &(229)\\ &= I^B &(230)\\ &= \left(I^B\right)^* &(231)\\ &= d\ \sum_z p_Z(z)|\phi_z^*\rangle\langle\phi_z^*| &(232)\\ &= \sum_z A_z^\dagger A_z. &(233)\end{aligned}$$

Now let us consider the action of the channel $\mathcal{M}$ on the maximally entangled state:

$$\mathcal{M}^{A\to B'}(\Phi^{BA}) \tag{234}$$

$$= \frac{1}{d}\sum_{z,i,j} |i\rangle\langle j|^B \otimes \sqrt{d\, p_Z(z)}|\psi_z\rangle\langle\phi_z^*||i\rangle\langle j||\phi_z^*\rangle\langle\psi_z|^{B'}\sqrt{d\, p_Z(z)} \tag{235}$$

$$= \sum_{z,i,j} p_Z(z)\ |i\rangle\langle j|^B \otimes \langle\phi_z^*|i\rangle\langle j|\phi_z^*\rangle\ |\psi_z\rangle\langle\psi_z|^{B'} \tag{236}$$

$$= \sum_{z,i,j} p_Z(z)\ |i\rangle\langle j|\phi_z^*\rangle\langle\phi_z^*|i\rangle\langle j|^B \otimes\ |\psi_z\rangle\langle\psi_z|^{B'} \tag{237}$$

$$= \sum_{z} p_Z(z)\ |\phi_z\rangle\langle\phi_z|^B \otimes\ |\psi_z\rangle\langle\psi_z|^{B'} \tag{238}$$

The last equality follows from recognizing $\sum_{i,j}|i\rangle\langle j|\cdot|i\rangle\langle j|$ as the transpose operation and noting that the transpose is equivalent to conjugation for an Hermitian operator $|\phi_z\rangle\langle\phi_z|$. Finally, since the action of both $\mathcal{N}_{\mathrm{EB}}^{A\to B'}$ and $\mathcal{M}^{A\to B'}$ on the maximally entangled state is the same, we can conclude that the two channels are equivalent (see Section 4.5). Thus, $\mathcal{M}$ is a representation of the channel with unit-rank Kraus operators. ∎

5 Summary

We give a brief summary of the main results in this lecture. We derived all of these results from the noiseless quantum theory and an ensemble viewpoint. An alternate viewpoint is to say that the density operator is the state of the system and then give the postulates of quantum mechanics in terms of the density operator. Regardless of which viewpoint you view as more fundamental, they are consistent with each other in standard quantum mechanics.

The density operator ρ for an ensemble $\{p_X(x), |\psi_x\rangle\}$ is the following expectation:

$$\rho = \sum_x p_X(x)|\psi_x\rangle\langle\psi_x|. \tag{239}$$

The evolution of the density operator according to a unitary operator U is

$$\rho \to U\rho U^\dagger. \tag{240}$$

A measurement of the state according to a measurement $\{M_j\}$ where $\sum_j M_j^\dagger M_j = I$ leads to the following post-measurement state:

$$\rho \to \frac{M_j\rho M_j^\dagger}{p_J(j)}, \tag{241}$$

where the probability $p_J(j)$ for obtaining outcome j is

$$p_J(j) = \mathrm{Tr}\left\{M_j^\dagger M_j\rho\right\}. \tag{242}$$

The most general noisy evolution that a quantum state can undergo is according to a completely-positive, trace-preserving map $\mathcal{N}(\rho)$ that we can write as follows:

$$\mathcal{N}(\rho) = \sum_j A_j\rho A_j^\dagger, \tag{243}$$

where $\sum_j A_j^\dagger A_j = I$.